Fuzzy Entropy-based State of Health Estimation for Li-Ion Batteries

Sui, Xin; He, Shan; Meng, Jinhao; Teodorescu, Remus; Stroe, Daniel-Ioan

Published in:
I E E E Journal of Emerging and Selected Topics in Power Electronics

DOI (link to publication from Publisher):
10.1109/JESTPE.2020.3047004

Publication date:
2021

Document Version
Accepted author manuscript, peer reviewed version

Fuzzy Entropy-based State of Health Estimation for Li-Ion Batteries

Xin Sui, *Student Member, IEEE,* Shan He, *Student Member, IEEE,* Jinhao Meng, *Member, IEEE*
Remus Teodorescu, *Fellow, IEEE,* Daniel-Ioan Stroe, *Member, IEEE*

***Abstract*—Accurate estimation of the state of health (SOH) of batteries is essential for maximizing the lifetime of the battery and improving the safety and economy of any energy storage system. Data-driven methods can use measurement data to effectively estimate the SOH, but the estimation performance depends on the relevance between the selected feature and SOH. In this paper, fuzzy entropy (FE) of battery voltage, is proposed as a new feature for SOH estimation and validated on Li-ion batteries. Compared with the traditional sample entropy, the FE can capture the variation of voltage during the battery degradation more efficiently in terms of the parameter selection, data noise, data size and test condition. Moreover, the aging temperature variation is involved in the established SOH estimator as the temperature is a disturbance variable in the real applications. The FE-SOH is used as the input-output data pair of the support vector machine, and a single-temperature model, a full-temperature model, and a partial-temperature model are established. As a result, the FE-based method has better estimation accuracy under aging temperature variation. The FE-based method also decreases the dependence on the size of the required training data. Finally, the effectiveness of the proposed method is verified by experimental results.**

***Index Terms*— Li-ion battery, state of health estimation, sample entropy, fuzzy entropy, aging temperature variation, support vector machine, short-term current pulse.**

I. Introduction

LI-ION batteries are extensively used in a wide range of applications including electric vehicles and energy storage systems [1]-[3]. However, similar to any energy storage device, their performance is subject to degradation (i.e., capacity fade and power decrease) during long-term operation. Hence, in order to ensure a reliable operation and economic viability, it becomes necessary to know the state of health (SOH) of the batteries at every time during their use [4]-[6]. SOH is a figure of merit of the condition of the battery. It is usually related to the battery capacity, resistance or to both, as these are the main parameters that are describing the battery performance behavior during their entire life. Various battery SOH estimation methods have been proposed [7]-[20]. The most straightforward approach is to measure the charge transferred through the battery during charging or discharging [7, 8]. However, in real applications, this method needs high-precision current sensors and requires the battery to be taken out of the regular operation, which is not feasible. Therefore, other estimation methods have to be considered. Based on electrochemical models or equivalent-circuit models, state observers such as multi-scale extended Kalman filter [9, 10], multi-scale nonlinear predictive filter [11], and the particle filter [12] have been designed for the joint estimation of battery state of charge (SOC) and SOH. Due to the complex internal principles and uncertain working conditions, it is difficult to establish an accurate battery model that can exhibit the battery dynamic characteristics. Data-driven methods are based on the battery aging features and evolution rules according to battery data (current, voltage, temperature). They are gaining increasing interest due to their flexibility and being battery model-free. These methods include amongst others, support vector machine [13, 14], relevance vector machine [15, 16], artificial neural network [17, 18], Gaussian process regression [19, 20], and ensemble learning [21], etc. Moreover, with the development of big data technology, real-time monitored parameters such as voltage, current, and temperature are processed in the cloud platform, which also decreases the requirements of the microcontrollers and improves the SOH estimation accuracy [5].

Because the data-driven SOH estimation methods model the battery degradation by mapping the external features against the capacity loss, high-quality datasets are required for the training purpose. Therefore, the feature should contain sufficient aging information to improve the estimation accuracy, and the convenience of feature extraction should be considered in practical applications. In [22], Saha et al. found that changes in the battery internal parameters such as the double layer capacitance and the charge transfer resistance are related to battery degradation, so they extract the features from the electrochemical impedance spectroscopy to estimate the SOH. However, such features are hard to be obtained in real-life applications due to the requirements and limitations of the measurement method and device. Based on the fully charge voltage profiles measured during the aging process, the incremental capacity (IC) peak, valley, and their corresponding voltage values [23, 24] can be extracted and related to the battery SOH. Nevertheless, obtaining the fully charging profile requires a long measurement time, which is not always feasible neither in laboratory research nor in practical applications. In

X. Sui, S. He, R. Teodorescu, and D. -I. Stroe are with the Department of Energy Technology, Aalborg University, Aalborg 9220, Denmark (e-mail: xin@et.aau.dk, she@et.aau.dk, ret@et.aau.dk, dis@et.aau.dk).

J. Meng is with the College of Electrical Engineering, Sichuan University, Chengdu 610065, China (e-mail: jinhao@scu.edu.cn).

order to shorten the test time for obtaining the features, Stroe et al. [25], and Xiong et al. [26] develop an effective feature for SOH estimation considering a partial charging voltage curve. In addition, the current pulses are added during the charging period of a battery. Based on the short-term pulse test, the direct current resistant [27], the geometric shape of the voltage profiles [28], and the knee points of voltage response to pulse test [14, 29] are extracted as the feature to estimate the SOH.

In addition, sample entropy (SE), as a powerful statistic for measuring the complexity of a signal, has been used for SOH estimation [30]-[33]. Li et al. monitored the surface temperature of batteries during the charging process, and utilized its SE for the battery capacity prediction [25]. Hu et al. calculated the SE of the voltage data under a short pulse test, and the SOH was estimated by the polynomial fitting [31] and the sparse Bayesian prediction [32]. Moreover, high accuracy of the entropy-based SOH estimation will be achieved when the battery SOC enters into the polarization zone [33].

However, SE shows high sensitivity to the parameter selection and noise because the Heaviside step function is used in its similarity degree computation [34]. Moreover, SE cannot reflect the information contained in the aging data accurately when the aging temperature is considered. In this paper, fuzzy entropy (FE) is applied for battery SOH estimation in order to improve the estimation accuracy. Because FE is an improved measure of time series regularity, FE of voltage will be more consistent and robust to parameter variation (i.e., test condition and date size) [35]. The advantages of FE-based method are verified from three aspects i.e., freer parameter selection, stronger robustness to noise, less dependence on data size, and more independent on aging test condition. Since the aging temperature has a significant effect on the battery's degradation behavior [36, 37], in this work, the performance of the proposed SOH estimation method for different temperature is studied in detail. According to the aging data under different aging temperature condition, the FE, SOH, and temperature are used as input variables. The FE-SOH mapping is established based on SVM, as shown in Fig. 1. Compared with the SE-based method, the effectiveness of the proposed method is verified by experiments.

The rest of this paper is organized as follows. The theory of SE and FE algorithms as well as the SVM-based method for SOH estimation are introduced in Section II. Section III compares the performance of FE-based and SE-based methods based on the cyclic aging test. Under the calendaring aging tests with different temperatures, the SOH estimation results are given in Section IV. Section V gives the conclusion of this work.

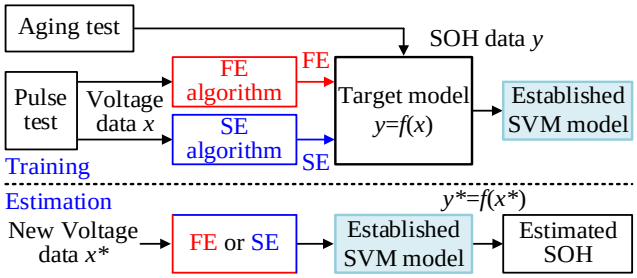


Fig. 1. Schematic diagram of the proposed algorithm.

II. Fuzzy Entropy-based SOH Estimation

A. The theory of FE and SE algorithm

SE or FE is the negative natural logarithm of the conditional probability (CP) that a dataset of length N, having repeated itself for m points within a boundary, will also repeat itself for m+1 points [34]. The specific algorithms of FE and SE are described in Fig. 2.

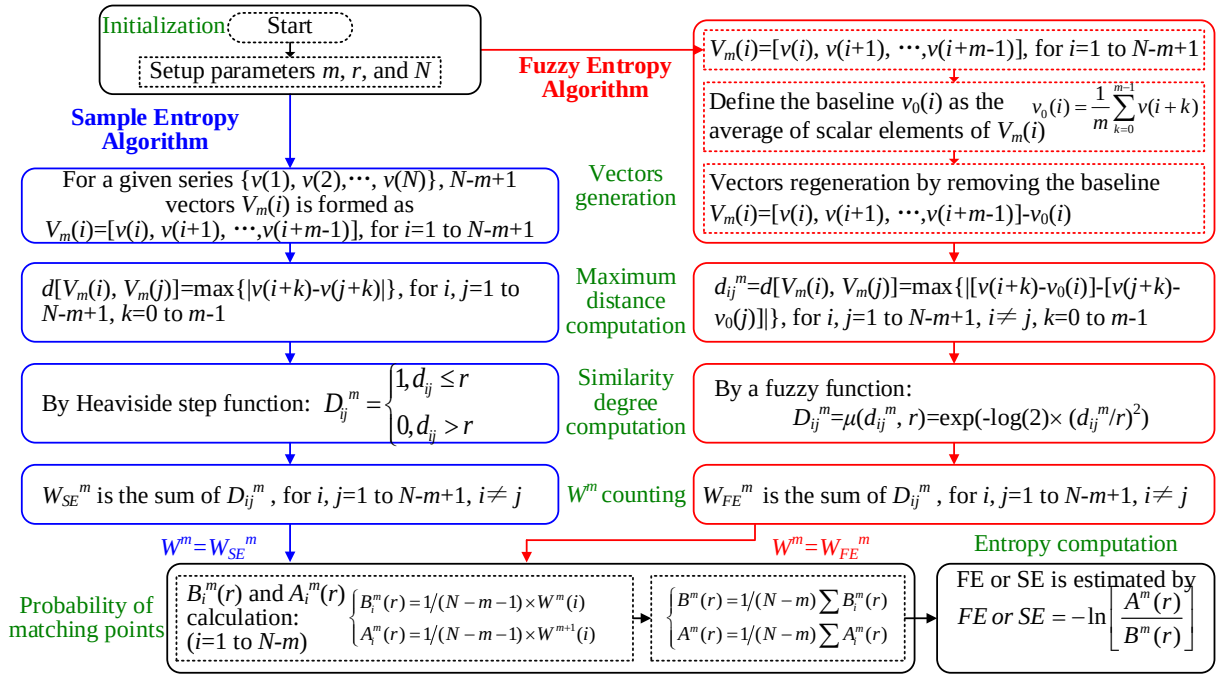


Fig. 2. Flowchart of fuzzy entropy and sample entropy algorithms.

There are two differences in the calculation of SE and FE, which makes FE more accurate in extracting the information contained in the original data than SE. Firstly, the similarity degree is computed by the Heaviside step function in SE; in this case, the contributions of two points that are far apart but within the boundary are treated equally, while the points just outside the boundary are left out. For example, there are four points d_1, d_2, d_3 and the original point d_0 in the series, as shown in Fig. 3. Because the Heaviside function has a rigid boundary, d_2 and d_3 have the same similarity degree with the original point d_0, i.e., $D_{d_2d_0} = D_{d_3d_0} = 1$. However, the point d_1, which is close to d_2, is considered dissimilar with d_0, just because d_1 is outside the boundary. As a result, a small variation of r will cause the saltation of SE value and SE maybe invalid in case of small parameter r [35]. FE improves the SE by using an exponential function as fuzzy function to describe the similarity degree of two vectors. In this case, the larger the distance between two points, the lower their similarity degree. Therefore, the fuzzy function can obtain more details of the data and it makes FE more accurate in quantifying the regularity of a dataset and freer in parameter selection. Secondly, the m-dimensional vector $V_m(i)$ is generated directly from the original series in SE. Then the distance between $V_m(i)$ and $V_m(j)$ is defined as the maximum absolute difference of their scalar elements. Under the definition, the similarity degree in SE is determined by the absolute coordinates of the vectors. However, the mean of the match templates is removed in the case of FE calculation, so that the similarity of the vectors is measured based on their shapes rather than their absolute coordinates. Therefore, FE can maintain a stronger relative consistency when there are fluctuations in the original series. From a theoretical perspective, FE has strong robustness to the data noise and the test condition. The good performance of FE is also demonstrated in the following sections.

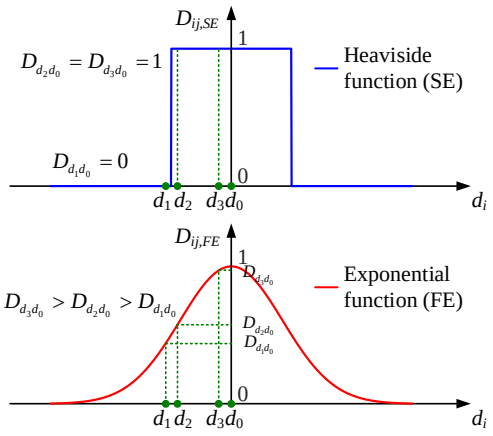


Fig. 3. The similarity function used to calculate SE and FE. The Heaviside function and the exponential function is for SE and FE calculation, respectively.

Typically, the parameter m is suggested to be set at 2 or 3, and r is to be set between 0.1 and 0.25 times the standard deviation of the data [35]. As these suggestions do not always demonstrate the best results for all kinds of datasets, the FE and SE algorithm can be tested using a range of parameter combinations (m = 2 and 3, r ranging from 0.1–0.3 times the standard deviation of the data) [38]. Then, the parameter can be chosen based on the minimization of the maximum sample/fuzzy entropy relative error [39]. In the strategy for the optimal selection of r, the standard approximation is used and its expression is

$$\sigma_{g(CP)} \cong \left|g'(CP)\right| \sigma_{CP} \tag{1}$$

where $g(CP)$=−log(CP). Then,

$$\sigma_{g(FE)} \cong \left|g'(FE)\right| \sigma_{CP} = \frac{\sigma_{CP}}{-\log(CP)CP} \tag{2}$$

$$\sigma_{g(CP)} \cong \left|g'(CP)\right| \sigma_{CP} = \frac{\sigma_{CP}}{CP} \tag{3}$$

where $\sigma_{g(FE)}$ and $\sigma_{g(CP)}$ are the relative errors of the FE and CP estimates, respectively. The parameter r can be selected by minimizing the quantity

$$\max\left(\frac{\sigma_{CP}}{-\log(CP)CP}, \frac{\sigma_{CP}}{CP}\right) \tag{4}$$

B. SOH estimation based on SVM

SVM is an effective method to deal with nonlinear regression problems, which uses kernel technique to map features vectors to high-dimensional space [14]. In this work, an SVM model is established to capture the nonlinear relationship between features (i.e., FE and SE) and SOH. In general, the SVM model is defined as

$$\hat{y} = \mathbf{w}^T \cdot \psi(\mathbf{x}) + b, \quad \mathbf{x} \in R^d, \ \psi(\mathbf{x}) \in R^{\tilde{d}}, \ b \in R \tag{5}$$

where $\mathbf{x}$ is the feature vector, y is the SOH value, and $\psi(\cdot)$ is the mapping that makes the input data linear in the new feature space. The ε-insensitive loss function, as expressed in (6), is used in the SVM model.

$$\xi_\varepsilon(\hat{y}, y) = \begin{cases} 0, & |y - \hat{y}| < \varepsilon \\ |y - \hat{y}| - \varepsilon, & \text{otherwise} \end{cases}, \ \forall i \in \{1, 2, \ldots, n\} \tag{6}$$

The objective of the SVM is to find the optimal coefficients $\mathbf{w}$ and b on the basis of the following constrained optimization problem,

$$\min_{\mathbf{w} \in R^d, \, \xi_i, \xi_i^* \in R^n} \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_i^n \left(\xi_i + \xi_i^*\right)$$
$$s.t. \begin{cases} y_i - \mathbf{w}^T \cdot \psi(\mathbf{x}_i) - b \le \varepsilon + \xi_i \\ \mathbf{w}^T \cdot \psi(\mathbf{x}_i) + b - y_i \le \varepsilon + \xi_i^*, \forall i \in \{1, 2, \ldots, n\} \\ \xi_i, \xi_i^* \ge 0 \end{cases} \tag{7}$$

where n is the sample number, C is a positive constant regulating the penalty, ξ_i and ξ_i^* are the slack variables creating a soft-margin. The Lagrangian function can be introduced to transfer such a problem into its dual problem which satisfies the Karushe-Kuhne-Tucker conditions [40]. Therefore, the following equations can be obtained.

$$\mathbf{w} = \sum_{i=1}^{n} \left(\alpha_i - \alpha_i^*\right) \psi(\mathbf{x}_i) \tag{8}$$

$$\sum_{i=1}^{n}\left(\alpha_i-\alpha_i^*\right)=0 \tag{9}$$

$$\begin{cases} C=\alpha_i+\xi_i \\ C=\alpha_i^*+\xi_i^* \end{cases} \tag{10}$$

$$b=y_i-\sum_{i=1}^{n}\left(\alpha_i-\alpha_i^*\right)\psi(\mathbf{x}_i)^T\cdot\psi(\mathbf{x}_i), \tag{11}$$

for examples i where $0<\alpha_i,\alpha_i^*<C$

Finally the regression function can be described as:

$$f(\mathbf{x})=\sum_{i=1}^{n}(\alpha_i^*-\alpha_i)^T K(\mathbf{x}_i,\mathbf{x})+b \tag{12}$$

where $K(\mathbf{x}_i, \mathbf{x})$ is the radial basis function kernel with the form of $K(\mathbf{x}_i,\mathbf{x})=\exp\left(-\|\mathbf{x}_i-\mathbf{x}\|^2/2\gamma\right)$, α_i^* and α_i are Lagrange multipliers. In order to avoid the overfitting problem, the optimal model parameters are obtained through 5-fold cross-validation [15]. The initial dataset denoted as D is randomly split into five mutually exclusive subsets D_1, D_2, …, D_5, and the model is trained and tested five times repeatedly. In each time, the model is trained on $D\backslash D_i$ and tested on D_i, where $i\in\{1, 2, …, 5\}$. The output of the cross-validation is the average of all the five testing errors.

III. Experimental Test and Aging Results

The main electrical parameters of the tested Li-ion battery cells are listed in Table 1. The experimental setup, as shown in Fig. 4, consists of a FuelCon programmable battery test station which is used to perform the reference measurements (i.e., capacity test and pulse test), and a host computer which is used for data acquisition and analysis.

TABLE I
THE DATASHEET OF THE LI-ION BATTERY

Item	Value
Chemistry	LFP/C
Nominal capacity	2.5 Ah
Nominal voltage	3.3 V
Maximum voltage	3.6 V
Cut-off voltage	2.0 V
Maximum continuous charge current	10 A
Maximum continuous discharge current	50 A

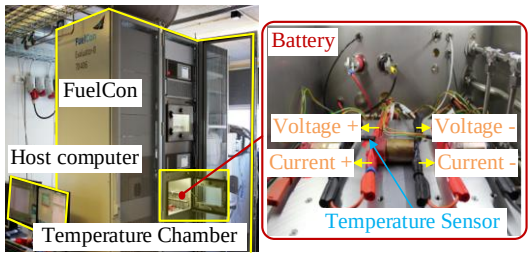


Fig. 4. Experimental setup.

During all the measurements, the battery cells were placed inside climatic chambers in order to reach and maintain the desired temperature. The temperatures mentioned in this paper, are the temperatures measured on the cell surface. Generally, battery aging can be divided into the calendar aging and cyclic aging. The calendar aging corresponds to the irreversible capacity loss caused by storage while the cyclic aging is associated with the impact of the charge/discharge cycles of the battery. In order to fully verify the effectiveness of the proposed FE feature, both aging dimensions were considered when aging the battery cells.

A. Cycle aging test

The whole test consists of three parts: the aging test, the capacity test, and the pulse test, as shown in Fig. 5. Firstly, the Li-ion battery was aged with a mission profile from the energy storage system providing primary frequency regulation to the grid [3]. Secondly, a capacity test was conducted to measure the battery capacity, as shown in Fig. 6(a). The capacity test consists of a 1C-rate constant current full charging procedure and a full discharging procedure. In this work, the battery SOH is calculated as the ratio between the current maximum available capacity and the initial maximum available capacity. Thirdly, using a 1C-rate constant current, the battery was charged to 20% SOC, 50% SOC and 80% SOC, respectively, and at each SOC level, a pulse test was conducted for FE/SE features extraction. Each pulse test lasted for 30s, including a 20 seconds of 4C-rate charging and a 10 seconds of relaxation. The obtained voltage datasets are shown in Fig. 6(b), where the behavior of the voltage curve changed during the degradation process of the battery. After the whole test, the battery reached its end-of-life criteria, which was predefined at 20% capacity fade. The aging mission profile had a length of one week, time in which the battery is subjected to approximately 600 Ah throughput (i.e., 120 full equivalent cycles (FEC)). The aging profile was repeated 38 times, resulting in a load of 4560 FECs for the tested battery.

B. Calendar aging test under various temperatures

The calendar aging test was carried out to simply analyze the effect of aging condition (in this case the aging temperature) on the performance of FE-based SOH estimation. The test process of calendar aging is similar with that of cyclic aging. As shown in Fig. 7, six battery cells, numbered C.1 to C.6, were aged at three different calendar conditions. These battery cells were charged to 50% SOC and they were stored at 55°C, 47.5°C, and 40°C, respectively. After each one-month calendar aging, the chamber was set to 25°C, then the capacity test and the pulse test were conducted. Each pulse test lasted for 33s, including a 10 seconds of 4C-rate charging, a 3 seconds of 4C-rate discharging, and two 10 seconds of relaxation. Fig. 8(a) and Fig. 8(b) show the measured SOH and the collected voltage datasets of training cells, respectively, from the calendar aging.

According to the obtained voltage data, the SE and FE can be extracted and the SOH estimation can be achieved by the SVM method. The data of the tested battery are divided into two parts: one part is used for SVM training, and the other part is used for verification. The root-mean-squared error (RMSE), the absolute percentage error (APE), and the mean absolute percentage error (MAPE) are used to evaluate the performance of the proposed method, which are defined as:

$$RMSE=\sqrt{\frac{1}{N_T}\sum_{i=1}^{N_T}\left(\hat{SOH}_i-SOH_i\right)^2} \tag{13}$$

$$APE_i = \frac{\left|\hat{SOH}_i - SOH_i\right|}{SOH_i} \times 100\% \quad (14)$$

$$MAPE = \frac{1}{N_T}\sum_{i=1}^{N_T}\left(\frac{\left|\hat{SOH}_i - SOH_i\right|}{SOH_i} \times 100\%\right) \quad (15)$$

where N_T is the total number of validation data, $\hat{SOH}_i$ and SOH_i is the estimated SOH and the real SOH of the ith validation data point, respectively.

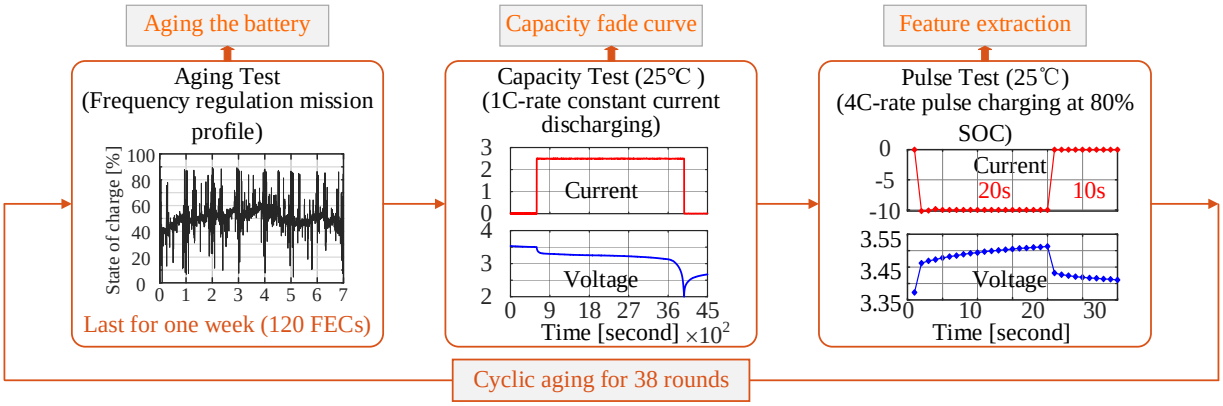
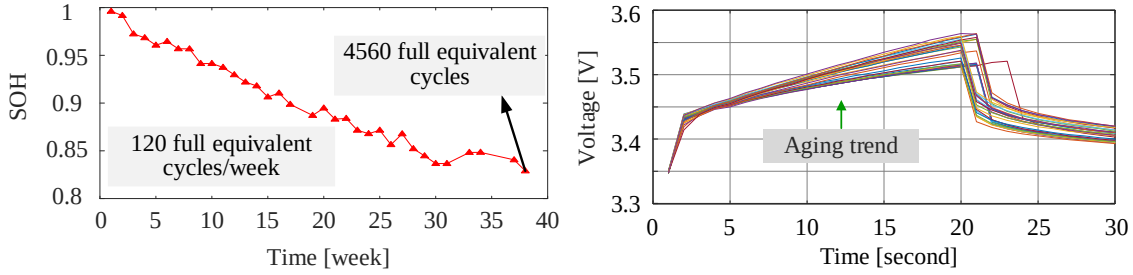


Fig. 5. Flowchart of the cyclic aging test schedules.



(a) (b)

Fig. 6. Experimental data obtained from the cyclic aging test. (a) SOH curve of the tested battery cell during the cyclic aging. (b)Voltage datasets obtained during the current pulse test.

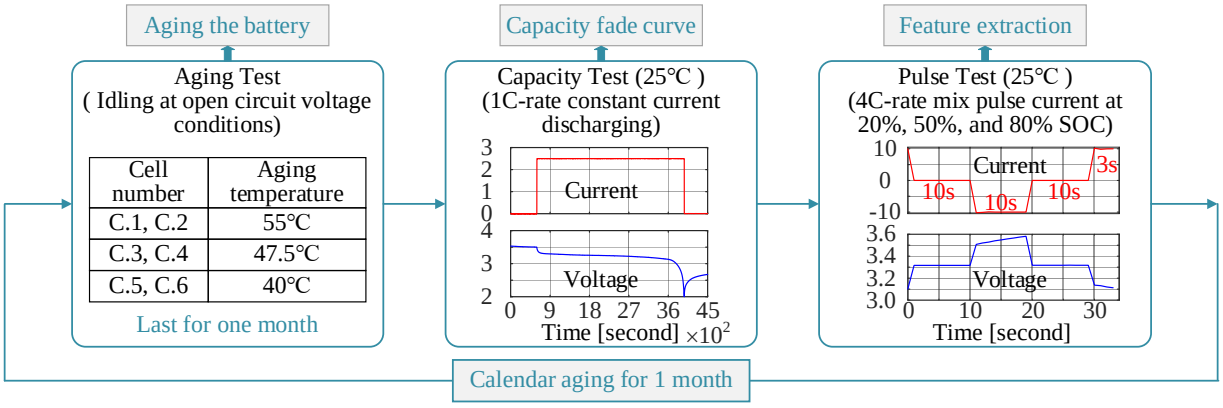
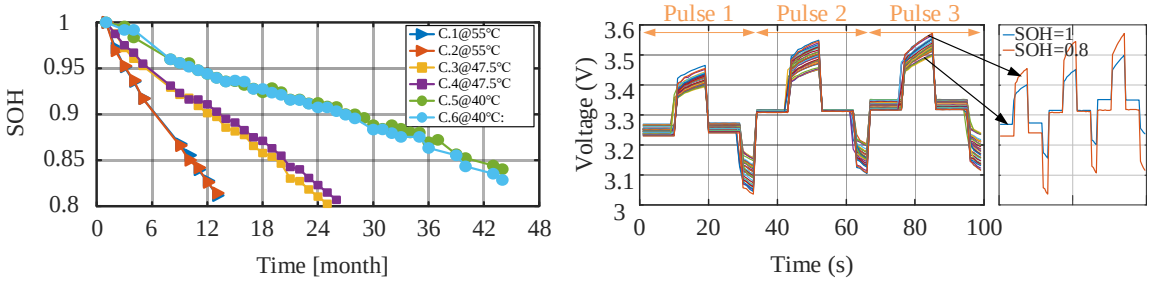


Fig. 7. Flowchart of the calendar aging test schedules.



(a) (b)

Fig. 8. Experimental data obtained from the calendar aging test. (a) SOH curves of the tested battery cells during the calendar aging, (b) Voltage datasets obtained during the current pulse tests.

IV. Performance Comparison of FE-based and SE-based Method

This section compares the characteristics of FE and SE when estimating the battery SOH in a dynamic operation condition. The 38 weeks of cyclic aging data are used to extract the features (SE and FE of voltage) and the SOH, which are used as input and output of the SVM model. Among these data, 28 feature-SOH pairs were used for model training and the other 10 data pairs were used for validation. Based on the experimental test, the effect of three factors on the estimation accuracy is considered respectively, i.e., the selection of parameter *r*, the noise and the size of the training data.

A. Effect of parameter selection on estimation accuracy

In terms of parameter selection, *m* is fixed to 2 and *N* is 30 (as there are total 30 voltage points in each pulse). Different values for *r* are selected to study its effect on the estimation accuracy. It can be seen from Fig. 9 that with the decreasing of *r* from 0.048 to 0.01, FE increases slightly and SE becomes scattered. Especially when *r* is set to 0.01, SE is invalid for SOH estimation. The SOH estimation results in terms of the variation of *r* are presented in Fig. 10(a) and Fig. 10(b). As shown in Fig. 10(a), both methods show high estimation accuracy when *r* is set to 0.048. However, according to Fig. 10(b), the estimation error of FE-based method is lower than 4% when *r* is 0.024, while the SOH estimated by the SE-based method shows large fluctuations.

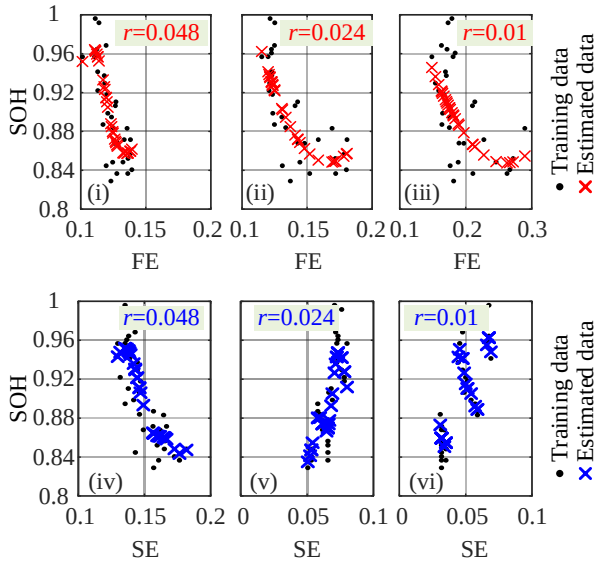


Fig. 9. SVM training results using different parameter *r*.

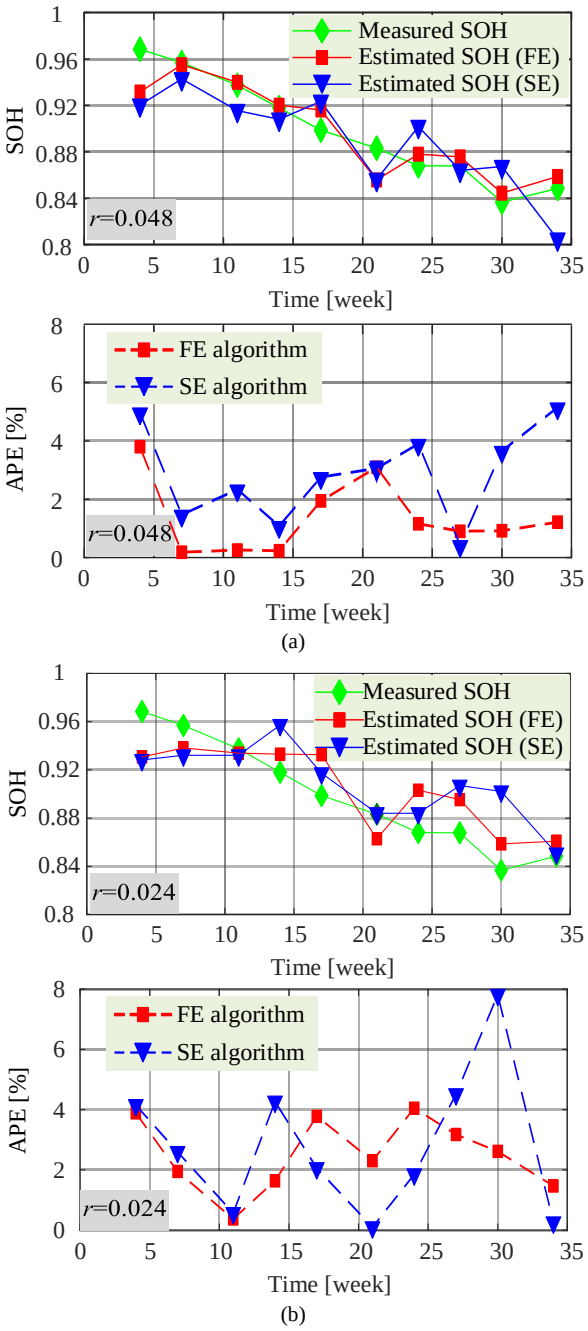


Fig. 10. Estimation resulrs with different paremeter *r*. (a) Estimation results when *r*=0.048, (b) Estimation results when *r*=0.024.

B. Effect of noise on estimation accuracy

In order to study the effect of noise on the estimation performance of FE-based and SE-based methods, Gaussian noise with a signal-to-noise ratio of 50 dB is added to the original voltage. The original voltage and noisy voltage are shown in Fig. 11. In this part, the parameter *m* and *r* are fixed to 2 and 0.048 for both methods, respectively. By calculating the entropy on the original voltage and noisy voltage, the variation of FE and SE can be seen in Fig. 12. One can observe that the monotonous relationship between FE and SOH is maintained when noise is added to the voltage signal. The noisy data causes the saltation of the SE, so the SE-based method becomes invalid for estimating the SOH. According to Fig. 13,

the FE-based method shows better accuracy and consistency when noisy data are used for SVM training and validation.

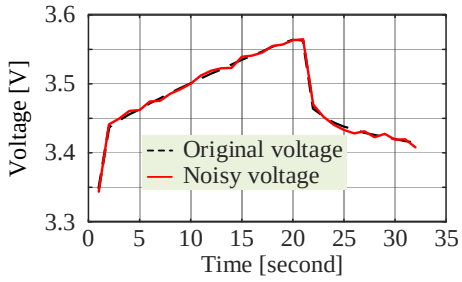


Fig. 11. Original voltage and noisy voltage.

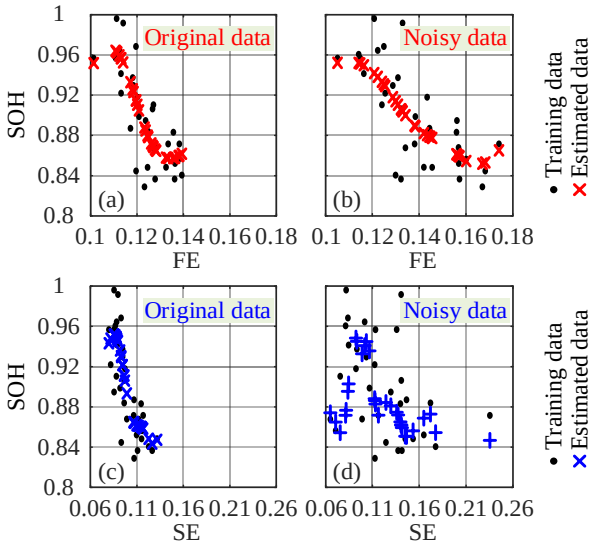


Fig. 12. Entropy value of original voltage and noisy voltage.

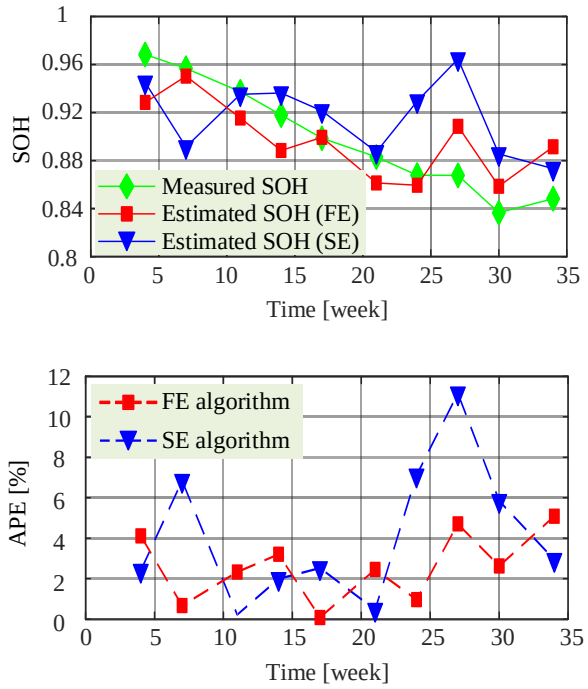


Fig. 13. Estimation results with Gaussian noise.

C. *Effect of data size on estimation accuracy*

The parameters *m*, *r* are still set at 2, 0.048 for FE and SE algorithms, and the number of data pairs for training varies from 16 to 29. As shown in Fig. 14(a), the FE-based method shows higher accuracy than SE-based method no matter how many data are used for SVM training. In addition, as shown in Fig. 14 (b), 2% MAPE can still be obtained for FE-based method using only approximately half of the data (16 FE-SOH data pairs), while SE-based method needs at least 24 SE-SOH data pairs to achieve the same error. Compared with the conventional SE-based method, the FE-based method can reduce the required data size effectively, which results in a less required effort for measurements either in laboratory or in real-life operation.

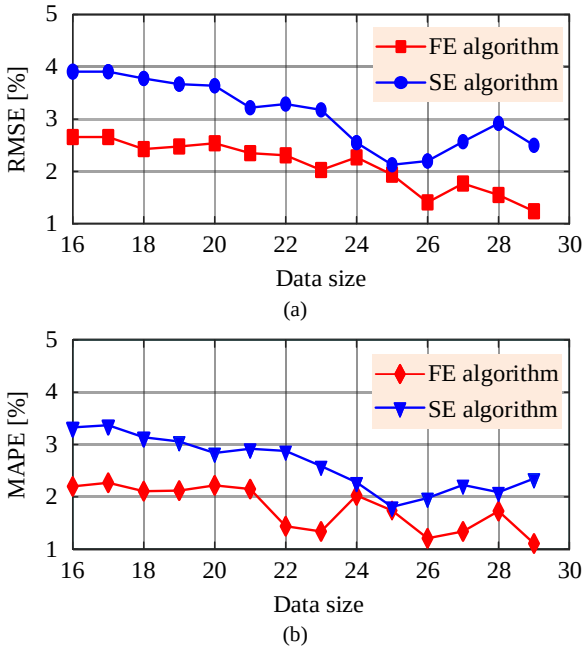


Fig. 14. Estimation results with different data size. (a) RMSE (b) MAPE.

D. *Effect of SOC on estimation accuracy*

In the previous sections (IV.A – IV.C), the SE and FE features were extracted from a voltage data set, which was measured at 80% battery SOC level (as seen in Fig. 6(b)). In this section, the effect of the SOC on the estimation accuracy of the battery SOH is analyzed. Thus, the SE and FE features were extracted for the three SOC levels and the obtained entropy values were related to the battery SOH.

As it can be observed in Fig. 15, SE obtained only for one SOC level cannot be used for predicting the battery SOH. On the other hand, FE feature obtained at a single SOC level can accurately predict the battery SOH. Furthermore, as illustrated in Fig. 15(b), the SOH estimation is less sensitive to the SOC at which the FE feature is obtained. This suggests that for the considered SOC levels (i.e., 20%, 50%, and 80% SOC), the battery SOH can be estimated accurately without the need of estimating correctly the SOC beforehand. Subsequently, the proposed feature and SOH estimation method is feasible in real-life applications as the battery system should not reach a certain SOC level before the 30 seconds current pulse is applied.

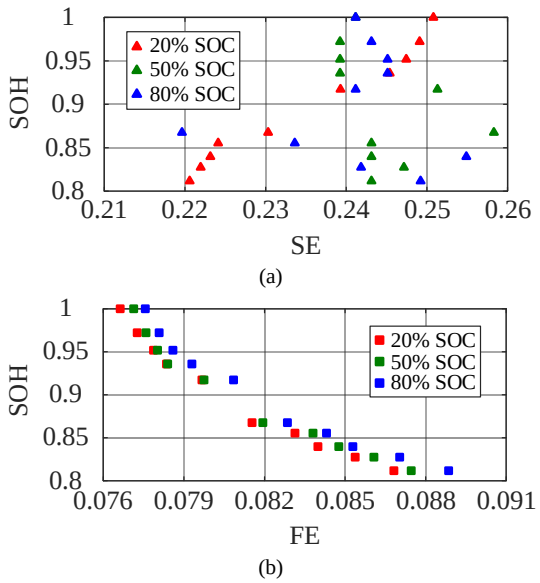


Fig. 15. Entropy values extracted from the voltage under single pulse test with different SOC levels. (a) SE, (b) FE. (Take the single pulse test at 47.5ºC for an example).

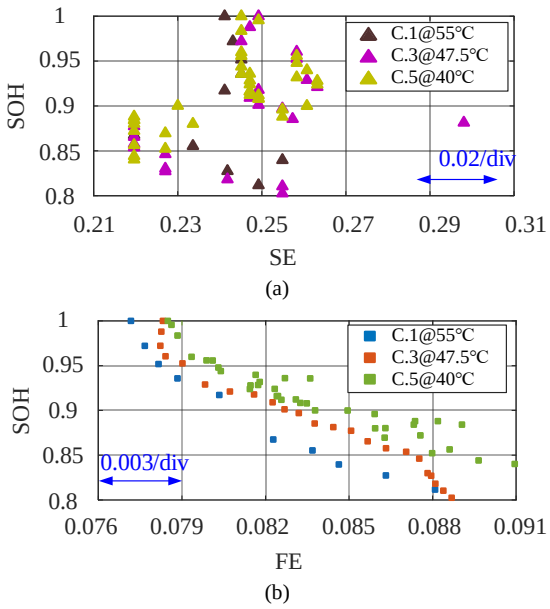


Fig. 16. Entropy values extracted from the voltage under single pulse test. (a) SE, (b) FE. (Take the single pulse test at 80% SOC for an example).

V. SOH ESTIMATION CONSIDERING TEMPERATURE VARIATION

The entropy feature may be invalid as the battery exhibits different degradation behavior when the aging temperature is varied. In order to avoid the effect of the current and SOC, and to simply investigate the effect of aging temperature on the SOH estimation accuracy, calendar aging tests under three temperatures were conducted. The temperature is added as an input variable and the coefficient of the SOH estimation model can be acquired by the SVM method.

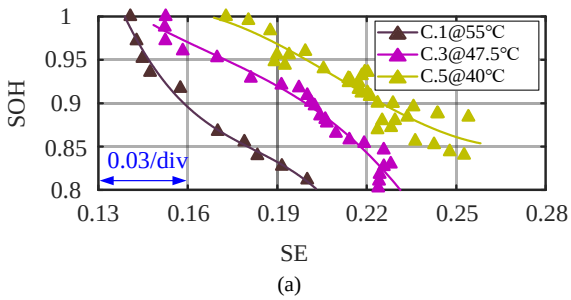
Before estimating the SOH considering the aging temperature, the paper first studies the minimum pulse time required for FE/SE calculation. As illustrated in Section III. B, a 33-second pulse test was conducted at 20% SOC, 50% SOC and 80% SOC, respectively (i.e., Pulse1, Pulse 2, and Pulse 3). It can be seen from Fig. 16 that when only the voltage data under one pulse is used for feature calculation, the obtained SE has no monotonous relationship with the SOH disappears. On the contrary, the obtained FE is still valid for SOH estimation. Here the single pulse test at 80% SOC for is taken as an example. Of course the same result can be also obtained for single pulse at other SOC levels (as seen in Fig. 15). Table II summarizes the training RMSE when using different pulse test data to calculate the FE and SE feature. It can be seen that 33-second pulse test (there are only 13 seconds of pulse current applied to the battery) is not enough for SE-based method to estimate the SOH. However, for FE-based method, its RMSE values keep lower than 0.015 when FE feature is obtained by using different amounts of voltage data in the range of 33 to 99 seconds. The results show that the pulse test containing at least 20-seconds pulse current is needed for SE-based method, while the FE-based method only requires a pulse test containing a 13-second pulse current.

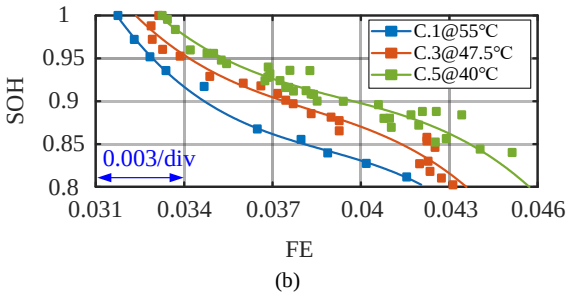
TABLE II
TRAINING RMSE WHEN USING DIFFERENT PULSE TEST DATA TO EXTRACT FE

Feature	Battery cell	Pulse 1	Pulse 2	Pulse 3	Pulse 1+Pulse 2	Pulse 1+Pulse 2+Pulse 3
FE	C. 1@55℃	0.011	0.015	0.011	0.013	0.011
	C. 3@47.5℃	0.015	0.009	0.009	0.012	0.014
	C. 5@40℃	0.013	0.012	0.012	0.012	0.012
SE	C. 1@55℃	/[a]	/	/	0.018	0.014
	C. 3@47.5℃	/	/	/	0.017	0.014
	C. 5@40℃	/	/	/	0.014	0.016

[a] The SE extracted by the data under the specific condition is invalid for SOH estimation

In order to compare the performance of SE-based and FE-based SOH estimation method, all the 99 voltage data are used for the feature calculation. Based on the minimization of the maximum FE or SE relative error, *m*, *r*, *N* were considered as 2, 0.04, 99 for the FE algorithm and 2, 0.08, 99 for the SE algorithm. As shown in Fig. 17(a) and 17(b), the SE/FE curves have the same shapes but shift to different degree along the horizontal axis when the aging temperature is changed. Fig. 17(a) shows that for a fixed SOH point, the distance between the SE values at adjacent temperature (i.e., 55ºC and 47.5ºC) is approximately 0.03; however, for the FE curves in Fig. 17(b), the distance is smaller than 0.003. Consequently, it can be concluded that FE is more robust to temperature variation than the SE.





(b)

Fig. 17. Entropy values extracted from the combination voltage under three single pulses. (a) SE, (b) FE.

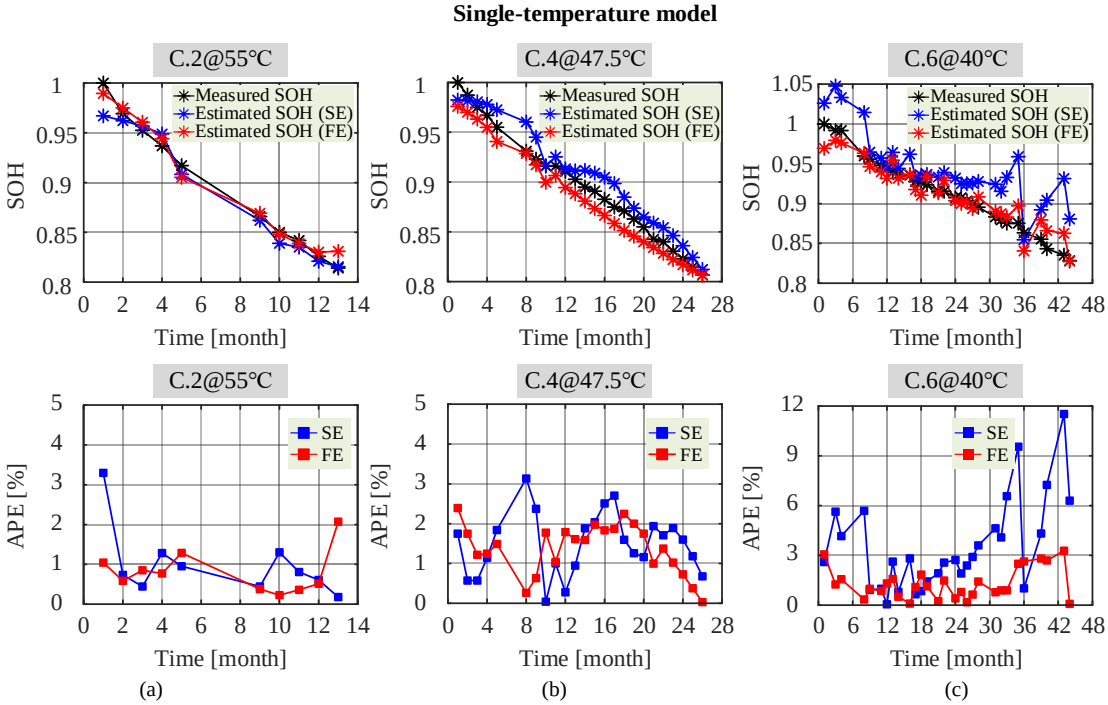
A. *Single-temperature model*

The mutual validation approach is used, where the battery cells are divided into two groups. One is the training group (C.1, C.3, and C.5) and the other is the validation group (C.2, C.4, and C.6). The data from only one temperature of aging test is used to train and validate the SVM model. Accordingly, three single-temperature models are established at 55℃, 47.5℃, and 40℃, respectively. As shown in Fig. 18, both methods can achieve an accurate SOH estimation and the maximum error is less than 4% at 55℃ and 47.5℃. While the SE-based method shows a big fluctuation and its APE even reaches 12% for the battery aged at 40℃. Meanwhile, FE-based method improve the estimation accuracy significantly, and the errors in most estimate values are less than 3%. Consequently, according to Fig. 19, the performance of both methods are approximately the same when the battery were aged at 55℃ and 47.5℃. However, SE-based method at 40℃ shows a large estimation error, and its RMSE and MAPE is 0.034 and 3.25%, respectively. On the contrary, FE-based method shows better estimated accuracy at 40℃, and the RMSE and MAPE are only about 0.5 times and 0.4 times smaller than that of SE-based method. It can be concluded that the FE-based method is more robust to the temperature variation than the SE-based method.

B. *Full-temperature model*

All the SE/FE-SOH data pairs at three temperatures are used for SVM training and a full-temperature model is established. In comparison to the SE-based method, as shown in Fig. 20 and Fig. 21, the FE-based method improves the estimation accuracy at 40℃. The RMSE of FE-based method decreases from 0.028 to 0.018, and the MAPE decreases from 2.44% to 1.39%.

In addition, the FE-based method is more robust to temperature variation than the SE-based method. According to Fig. 20, for all cases of validation, the APE of the FE-based method is consistently less than 3% (except for the 46th month estimate value of C.6 at 40℃). On the other hand, the errors at some points are larger than 3% for the SE-based method and the maximum APE is about 8% especially for C.6, which was aged at 40℃. Accordingly, one can observe from Fig. 21 that the RMSE of FE-based method changes slightly from 0.012 to 0.018, and the MAPE changes from 0.96% to 1.46%. It is worth noting that both the RMSE and the MAPE of the SE-based method show large fluctuation over a wide range temperature. That is because in the established SE-based model, there is a tendency towards data fitting at the intermediate temperature. As a result, for the SE-based method at 47.5℃, the RMSE and MAPE reach the smallest values which are only 0.011 and 1.12%. However, both errors increase obviously at the other two temperatures, especially for C.6 at 40℃, the RMSE and MAPE is 0.028 and 2.44%, respectively. It can be concluded that the FE-based method is more robust to the temperature variation than the SE-based method, and it will produce a relatively small estimation error independent on the battery aging temperature.



(a) (b) (c)

Fig. 18. Estimation results of single-temperature models. (a) C.2 at 55℃, (b) C.4 at 47.5℃, (c) C.6 at 40℃. (Each model is only trained with aging data at one temperature, and verified with the aging data of another battery at the same temperature).

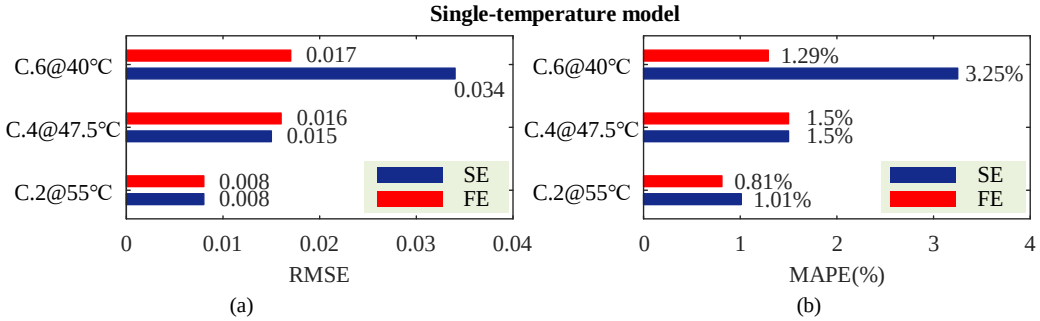


Fig. 19. Estimation errors of single-temperature models. (a) RMSE, (b) MAPE.

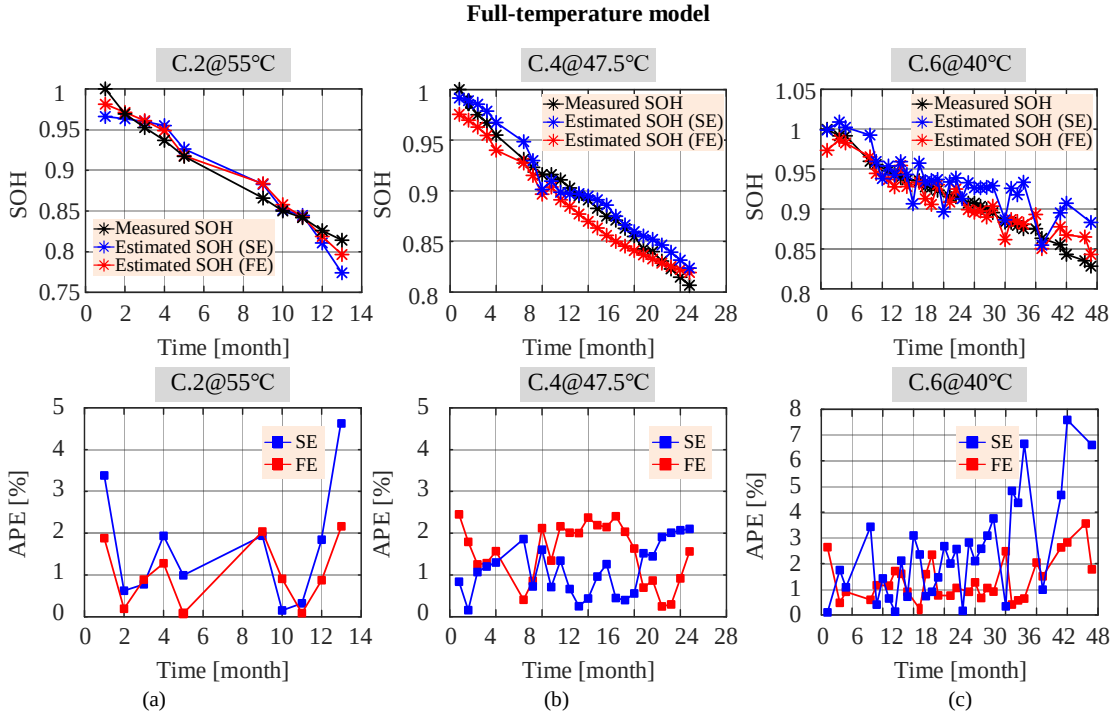


Fig. 20. Estimation results of full-temperature model. (a) C.2 at 55℃, (b) C.4 at 47.5℃, (c) C.6 at 40℃. (The model is trained using all data from C.1@55℃, C.3@47.5℃, and C.5@40℃, while the validations are performed separately for the aging data at each temperature).

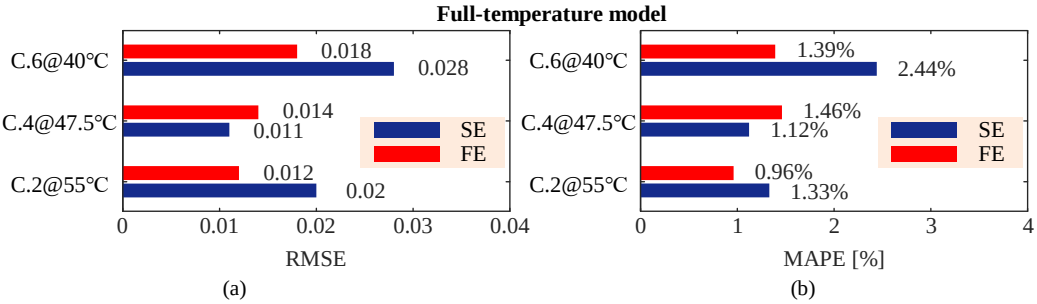


Fig. 21. Estimation errors of full-temperature model. (a) RMSE, (b) MAPE.

C. Partial-temperature model

In order to illustrate the advantage of FE-based method in the dependence of data size, the partial-temperature model is established by reducing the input temperatures from three to two. The training batteries (i.e., C.1 and C.5) are aged in the highest and the lowest temperature, and the validation battery C.4 was aged at the intermediate temperature. In comparison with the SE-based method, as shown in Fig. 22, the estimation errors of FE-based method are smaller than that of SE-base method in most of the time range. Moreover, according to Fig. 23, the RMSE and MAPE of FE-based method decrease from 0.012 to 0.011 and from 1.15% to 0.9%, respectively. Hence the FE-based method has better estimation accuracy than the SE-based method with less data, which illustrates that the FE-based method relies on less training data than the SE-based method.

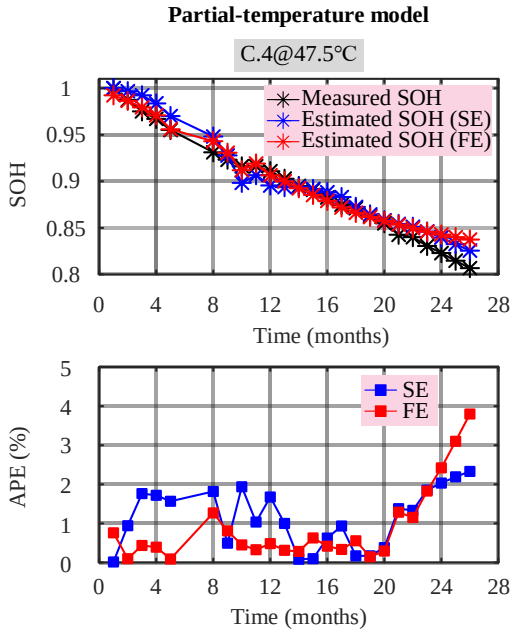


Fig. 22. Estimation results of partial-temperature model. (Training battery are C.1@55℃ and C.5@40℃, validation battery is C.4@47.5℃).

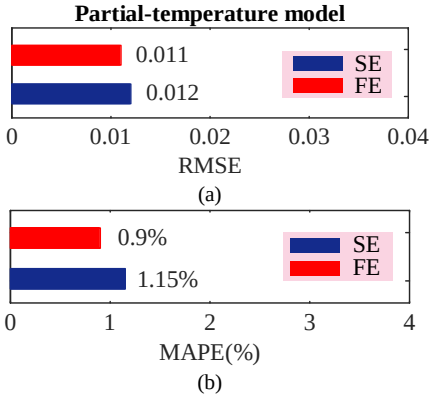


Fig. 23. Estimation errors of partial-temperature model for C.4. (a) RMSE, (b) MAPE.

VI. Conclusions

This paper proposes the FE of voltage as a new feature for Li-ion battery SOH estimation, and the nonlinear relationship between the SOH and the FE is established by SVM. A cyclic aging test is used to assess the performance of both SE- and FE-based methods. The results show that FE is more suitable than SE to estimate the SOH in three aspects. Firstly, FE is more stable with the variation of parameter *r*. Secondly, FE shows better consistency when noise is added to the voltage data. Thirdly, the FE-based method has higher accuracy when reducing the training data size. Fourthly, FE is more independent on the aging condition (i.e., the SOC level and the aging temperature).

In order to further validate the advantages of the FE-based method over the SE-based method, the performance for both methods in terms of the temperature variation is considered. Using the aging temperature as the input variable, the single-temperature model, full-temperature model, and partial-temperature model are established based on data obtained from three calendar aging tests. For the single-temperature model, the SE-based method shows a large estimation error while the FE-based method still has a high estimation accuracy. This is because the temperature variation in the training data can be regarded as noise, and SE is sensitive to the noise. For the full-temperature model, the model for the SE-based method has a tendency towards fitting data at one specific temperature. Consequently, the estimation error will be large in the other temperatures and will only be small at a specific temperature. On the contrary, the FE-based method always shows a relatively high estimation accuracy independent on the aging temperature. For the partial-temperature model, the data at the lowest and the highest temperature are used to estimate the SOH at the intermediate temperature. The FE-based method again shows a higher estimation accuracy than the SE-based method (i.e., 0.9% MAPE for the FE-based method and 1.15% MAPE for the SE-based method), which suggests that the proposed FE-based method is more robust to the data size than the traditional SE-based method.

References

[1] M. Berecibar, I. Gandiaga, I. Villarreal, N. Omar, J. Van Mierlo, and P. Van den Bossche, "Critical review of state of health estimation methods of Li-ion batteries for real applications," *Renew. Sust. Energ. Rev.*, vol. 56, pp. 572-587, April, 2016.

[2] D. I Stroe, M. Swierczynski, A. Stroe, R. Laerke, P.C. Kjaer and R. Teodorescu, "Degradation behavior of lithium-ion batteries based on lifetime models and field measured frequency regulation mission profile," *IEEE Trans. Ind. Appl.*, vol. 52, no. 6, pp. 5009-5018, Nov. 2016.

[3] D. I. Stroe, V. Knap, M. Swierczynski, A.I. Stroe, and R. Teodorescu, "Operation of a grid-connected lithium-ion battery energy storage system for primary frequency regulation: A battery lifetime perspective," *IEEE Trans. Ind. Appl.*, vol. 53, no. 1, pp. 430-438, Jan. 2017.

[4] X. Hu, F. Feng, K. Liu, L. Zhang, J. Xie, B. Liu, "State estimation for advanced battery management: key challenges and future trends," *Renew. Sust. Energ. Rev.*, vol. 114, pp. 109334, 2019.

[5] R. Xiong, L. Li, and J. Tian, "Towards a smarter battery management system: A critical review on battery state of health monitoring methods," *J. Power Sources*, vol. 405, pp. 18–29, Nov. 2018.

[6] Y. Li, K. Liu, A.M. Foley, A. Zülke, M. Berecibar, E. Nanini-Maury, J. Van Mierlo, and H.E. Hoster, "Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review," *Renew. Sust. Energ. Rev.*, vol. 113, pp. 109254, Oct. 2019.

[7] W. Waag, S. Käbitz, D. U. Sauer, "Experimental investigation of the lithium-ion battery impedance characteristic at various conditions and aging states and its influence on the application," *Appl. Energy*, vol. 1, no. 102, pp. 885-97, Feb. 2013.

[8] D. I. Stroe, M. Swierczynski, A. Stroe, S. K. Kaer and R. Teodorescu, "Lithium-ion battery power degradation modelling by electrochemical impedance spectroscopy," *IET Renew. Power Gener.*, vol. 11, no. 9, pp. 1136-1141, Jun. 2017.

[9] K. S. Ng, C.S. Moo, Y.P. Chen, and Y.C. Hsieh, "Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries," *Appl. Energy*, vol. 86, no. 9, pp. 1506–1511, Sep. 2009.

[10] C. Zou, C. Manzie, D. Neši´c, and A.G Kallapur, "Multi-time-scale observer design for state-of-charge and state-of-health of a lithium-ion battery," *J. Power Sources*, vol. 335, pp. 121–130, Dec. 2016.

[11] Y. Hua, A. Cordoba-Arenas, N. Warner, and G. Rizzoni, "A multi time-scale state-of-charge and state-of-health estimation framework using nonlinear predictive filter for lithium-ion battery pack with passive balance control," *J. Power Sources*, vol. 280, pp. 293–312, Apr. 2015.

[12] J. Wei, G. Dong, and Z. Chen, "Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5634-5643, Dec. 2017.

[13] V. Klass, M. Behm, and G. Lindbergh, "A support vector machine-based state-of-health estimation method for lithium-ion batteries under electric

vehicle operation," *J. Power Sources*, vol. 270, no. 15, pp. 262-272, Dec. 2014.

[14] J. Meng, L. Cai, G. Luo, D.I. Stroe, and R. Teodorescu, "Lithium-ion battery state of health estimation with short-term current pulse test and support vector machine," *Microelectron. Reliab.*, vol. 88-90, pp. 1216–1220, Sep. 2018.

[15] J. Meng, L. Cai, D.I. Stroe, J. Ma, G. Luo, and R. Teodorescu, "An optimized ensemble learning framework for lithium-ion battery state of health estimation in energy storage system," *Energy*, vol. 206, pp. 118140, Sep. 2020.

[16] H. Li, D. Pan and C. L. P. Chen, "Intelligent prognostics for battery health monitoring using the mean entropy and relevance vector machine," *IEEE Trans. Syst., Man, Cybern. Syst*., vol. 44, no. 7, pp. 851-862, Jul. 2014.

[17] Y. Zhou, M. Huang, Y. Chen, and Y. Tao, "A novel health indicator for on-line lithium-ion batteries remaining useful life prediction," *J. Power Sources*, vol. 321, pp, 1-10, Jul. 2016.

[18] H. Dai, G. Zhao, M. Lin, J. Wu and G. Zheng, "A novel estimation method for the state of health of lithium-ion battery using prior knowledge-based neural network and Markov chain," *IEEE Trans. Ind. Electron*., vol. 66, no. 10, pp. 7706-7716, Oct. 2019.

[19] Y. Che, Y. Liu, Z. Cheng and J. Zhang, "SOC and SOH identification method of Li-ion battery based on SWPSO-DRNN," *IEEE J. Emerg. Sel. Topics Power Electron*., *Early Access*, 2020.

[20] P. Tagade, K.S. Hariharan, S. Ramachandran, A. Khandelwal, A. Naha, S.M. Kolake, and S.H. Han, "Deep Gaussian process regression for lithium-ion battery health prognosis and degradation mode diagnosis," *J. Power Sources*, vol. 445, pp. 227281, Jan. 2020.

[21] K. Liu, X. Hu, Z. Wei, Y. Li and Y. Jiang, "Modified Gaussian process regression models for cyclic capacity prediction of lithium-ion batteries," *IEEE Trans. Transport. Electrific*., vol. 5, no. 4, pp. 1225-1236, Dec. 2019.

[22] B. Saha, K. Goebel, S. Poll and J. Christophersen, "Prognostics methods for battery health monitoring using a Bayesian framework," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 2, pp. 291-296, Feb. 2009.

[23] D. Stroe and E. Schaltz, "Lithium-Ion battery state-of-health estimation using the incremental capacity analysis technique," *IEEE Trans. Ind. Appl*., vol. 56, no. 1, pp. 678-685, Jan.-Feb. 2020.

[24] Y. Li, M. Abdel-Monem, R. Gopalakrishnan, M. Berecibar, E. Nanini-Maury, N. Omar, P. van den Bossche, and J. Van Mierlo, "A quick on-line state of health estimation method for Li-ion battery with incremental capacity curves processed by Gaussian filter," *J. Power Sources*, vol. 373, pp. 40-53, Jan. 2018.

[25] D. I., Stroe, V. Knap, and E. Schaltz, "State-of-health estimation of lithium-ion batteries based on partial charging voltage profiles," *ECS Trans*., vol. 85, no. 13, pp. 379, May, 2018.

[26] R. Xiong, Y. Zhang, J. Wang, H. He, S. Peng and M. Pecht, "Lithium-ion battery health prognosis based on a real battery management system used in electric vehicles," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4110-4121, May 2019.

[27] K. M. Tsang and W. L. Chan, "State of health detection for Lithium ion batteries in photovoltaic system," *Energy Convers. Manag*., vol. 65, pp. 7-12, Jan. 2013.

[28] J. Wu, Y. Wang, X. Zhang, and Z. Chen, "A novel state of health estimation method of Li-ion battery using group method of data handling," *J. Power Sources*, vol. 327, pp. 457-464, Sep. 2016.

[29] C. Lei, J. Meng, D. Stroe, J. Peng, G. Luo and R. Teodorescu, "Multi-objective optimization of data-driven model for lithium-ion battery SOH estimation with short-term feature," *IEEE Trans. Power Electron*., pp. 1-1, Apr. 2020.

[30] J. Li, C. Lyu, L. Wang, L. Zhang, and C. Li, "Remaining capacity estimation of Li-ion batteries based on temperature sample entropy and particle filter," *J. Power Sources*, vol. 268, pp. 895-903, Dec. 2014.

[31] X. Hu, S. E. Li, Z. Jia, and B. Egardt, "Enhanced sample entropy based health management of Li-ion battery for electrified vehicles," *Energy*, vol. 64, pp. 953-960, Jan. 2014.

[32] X. Hu, J. Jiang, D. Cao, and B. Egardt, "Battery health prognosis for electric vehicles using sample entropy and sparse Bayesian predictive modeling," *IEEE Trans. Ind. Electron*., vol. 63, no. 4, pp. 2645–2656, Apr. 2016.

[33] X. Sui, D. I. Stroe, S. He, X. Huang, J. Meng, and R. Teodorescu, "The effect of voltage dataset selection on the accuracy of entropy-based capacity estimation methods for lithium-ion batteries," *Appl. Sci*., vol. 9, no. 19, pp. 4170, Oct. 2019.

[34] J. S. Richman, J. R Moorman, "Physiological time series analysis using approximate entropy and sample entropy," *Am. J. Physiol-Heart C*., vol. 278, no. 6, pp. H2039–H2049, Jun. 2000.

[35] W. Chen, Z. Wang, H. Xie and W. Yu, "Characterization of surface EMG signal based on fuzzy entropy," *IEEE Trans. Neural Syst. Rehabil. Eng*., vol. 15, no. 2, pp. 266-272, June 2007.

[36] F. Leng, C. M. Tan, and M. Pecht, "Effect of temperature on the aging rate of Li ion battery operating above room temperature," *Sci. Rep*., no. 5, pp. 12967, Aug.2015.

[37] X. Hu, L. Xu, X. Lin, M. Pecht, "Battery lifetime prognostics," *Joule*, vol. 4, no. 2, pp. 310-346, 2020.

[38] J. M. Yentes, N. Hunt, K. K. Schmid, J. P. Kaipust, D. McGrath, and N. Stergiou, "The appropriate use of approximate entropy and sample entropy with short data sets," *Ann. Biomed. Eng*., vol. 41, no. 2, pp. 349–365, Feb. 2013.

[39] D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman, "Sample entropy analysis of neonatal heart rate variability," *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, vol. 283, no. 3, pp. R789–R797, Sep. 2002.

[40] H.W. Kuhn, and A.W. Tucker, "Nonlinear programming," *in Proc. Berkeley Symp. Math. Statist. Probab*., 1951, pp. 481-492.

Xin Sui (S'17) received the B.Eng. degree in Electrical Engineering from Northeast Electric Power University, Jilin, China, in 2015, and the M.Sc. degree in Electrical Engineering from Institute of Electrical Engineering, Chinese Academy of Sciences, Beijing, China, in 2018. She is currently working toward the Ph.D. degree in Electrical Engineering at Aalborg University, Aalborg, Denmark. Her research interests include battery state of health estimation, lifetime prediction based on machine learning algorithm.

Shan He (S'17) received the B.Eng. degree in Electrical Engineering from Northeast Electric Power University, Jilin, China, in 2015, and the M.Sc. degree in Electrical Engineering from Zhejiang University, Hangzhou, China, in 2018. He is supported by the China Scholarship Council and currently working toward the Ph.D. degree in Electrical Engineering at Aalborg University, Aalborg, Denmark. His research interests include multiphase induction motor and grid-connected inverter.

Jinhao Meng (M'20) received the M.S. degree in control theory and control engineering and the Ph.D. degree in electrical engineering from Northwestern Polytechnical University, Xi'an, China, in 2013 and 2019, respectively. He was supported by the China Scholarship Council as a joint Ph.D. student with the Department of Energy Technology, Aalborg University, Aalborg, Denmark. He is currently an associate researcher in Sichuan University, Chengdu, China. His research interests include battery modeling, battery states estimation, and energy management of battery energy storage system.

Remus Teodorescu (F'12) received the Dipl.Ing. degree in electrical engineering from the Polytechnical University of Bucharest, Bucharest, Romania, in 1989, and the Ph.D. degree in power electronics from the University of Galati, Galati, Romania, in 1994. In 1998, he joined the Power Electronics Section, Department of Energy Technology, Aalborg University, Aalborg, Denmark, where he is currently a Full Professor. Since 2013, he has been a Visiting Professor with Chalmers University. His research interests include design and control of grid-connected converters for photovoltaic and wind power systems, high voltage dc/flexible ac transmission systems based on modular multilevel converters, and storage systems based on Li-ion battery technology including modular converters and active battery management systems.

Daniel-Ioan Stroe (M'11) received the Dipl.Ing. degree in automatics from the Transilvania University of Brasov, Brasov, Romania, in 2008, and the M.Sc. degree in wind power systems and the Ph.D. degree in lifetime modelling of Lithium-ion batteries from Aalborg University, Aalborg, Denmark, in 2010 and 2014, respectively. He is currently an Assistant Professor with the Department of Energy Technology, Aalborg University. He was a Visiting Researcher at RWTH Aachen, Germany, in 2013. He has co-authored more than 70 journals and conference papers. His current research interests include energy storage systems for grid and e-mobility, Lithium-based batteries testing and modelling, and lifetime estimation of Lithium-ion batteries.